DOCUMENT RESUME

ED 373 463 EC 303 231

AUTHOR	Gonzalez, Patricia
TITLE	Effectiveness, Compliance, Accountability, and Program Improvement: Perspectives on the Evaluation Requirement. Final Report.
INSTITUTION	National Association of State Directors of Special Education, Alexandria, VA.
SPONS AGENCY	Special Education Programs (ED/OSERS), Washington, DC.
PUB DATE	28 Jun 94
CONTRACT	HS92015001
NOTE	18p.; Prepared by Project FORUM.
PUB TYPE	Viewpoints (Opinion/Position Papers, Essays, etc.) (120)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Accountability; *Compliance (Legal); *Disabilities; *Educational Assessment; Educational Legislation; Elementary Secondary Education; Evaluation Methods; Evaluation Utilization; Federal Aid; Federal Legislation; Federal Programs; Individualized Education Programs; Outcomes of Education; *Program Effectiveness; *Program Evaluation; Program Improvement; Special Education
IDENTIFIERS	*Individuals with Disabilities Education Act Part B

ABSTRACT

This paper discusses program evaluation as a condition for the receipt of federal funds under Part B of the Individuals with Disabilities Education Act. It focuses on some actual and potential linkages between effectiveness or outcome evaluation, implementation evaluation, compliance monitoring, and the concepts of program improvement and accountability. A brief discussion of the utilization of evaluation results for accountability and other purposes is presented. The paper proposes an evaluation system that includes both compliance monitoring and program evaluation. Recommendations for reevaluation of the annual evaluation requirement include: (1) state education agencies need to determine the relative value of both information that permits a judgement of effectiveness and information that suggests improvements in implementation, particularly with respect to evaluating individualized education programs (IEPs); (2) decisions at the federal and state levels need to be made regarding the overall purpose of program evaluation in special education; (3) the marked imbalance between the resources devoted to compliance monitoring and program evaluation needs to be reconsidered; and (4) states without an effectiveness and implementation evaluation system might look to approaches in which local education agencies provide annual evaluative information on a few priority questions each year, including the IEPs. (Contains 14 references.) (JDD)

ED 373 463

EFFECTIVENESS, COMPLIANCE, ACCOUNTABILITY, AND PROGRAM IMPROVEMENT: PERSPECTIVES ON THE EVALUATION REQUIREMENT

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it
☐ Minor changes have been made to improve reproduction quality
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

Patricia Gonzalez, Ph.D.

Final Report
Year 1 Deliverable #6-1-4
Under Contract No. HS92015001

June 28, 1994

Prepared by:
Project FORUM
Joy Hicks, Project Director

Prepared for:
Office of Special Education Programs
U.S. Department of Education

National Association of State Directors of Special Education
1800 Diagonal Road, Suite 320
Alexandria, VA 22314

EC 303231

BEST COPY AVAILABLE

This report was supported in whole or in part by the U.S. Department of Education (Contract No. HS92015001). However, the opinions expressed herein do not necessarily reflect the position or policy of the U.S. Department of Education, and no official endorsement by the Department should be inferred.

ACKNOWLEDGEMENTS

The staff of Project FORUM would like to extend their sincere appreciation to the following individuals whose reviews served to enrich the quality and accuracy of this document.

Dr. Marsha Brauen
Project Director, Westat Corporation
Rockville, MD

Dr. Lizanne DeStefano
Assistant Professor, Educational Psychology
University of Illinois - Urbana/Champaign
Champaign, IL

Mr. Wayne Erickson
Manager, Special Education Section
Minnesota State Board of Education
St. Paul, MN

Dr. Gerrie Hawkins
Program Review Branch
Office of Special Education Programs
U.S. Department of Education

TABLE OF CONTENTS

Abstract iii

The Big Picture: Evaluating Effectiveness and Implementation 2
Implications of Effectiveness Evaluation 2
Implications of Implementation Evaluation 3

Defining the Purpose of Evaluation 5
A System for Program Improvement 5
A System for Accountability 6

Influencing the Utility of Evaluation Results 8

Reconceptualizing the Annual Evaluation Requirement 9

References 11

ABSTRACT

This paper discusses some actual and potential linkages between effectiveness or outcome evaluation, implementation evaluation, compliance monitoring, and the concept of accountability. In light of this discussion, the paper proposes an evaluation system that includes both compliance monitoring and program evaluation. A brief discussion of the utilization of evaluation results for accountability and other purposes is also presented.

EFFECTIVENESS, COMPLIANCE, ACCOUNTABILITY, AND PROGRAM IMPROVEMENT: PERSPECTIVES ON THE EVALUATION REQUIREMENT

Program evaluation as a condition for the receipt of Federal funds under Part B of the Individuals with Disabilities Education Act (originally titled the Education for All Handicapped Children Act [EHA]) was intended to ensure implementation of program requirements, as well as compliance with the major provisions of the Act. At the time the EHA was written, it was felt that appropriate implementation of the individual education program (IEP) would be a strong indicator of compliance with the overall Act (W. Halloran, personal communication, November 24, 1993).

Currently, Part B of the Individuals with Disabilities Education Act (IDEA) contains the same wording as the original act: each State Plan is to "provide for procedures for evaluation at least annually of the effectiveness of programs in meeting the educational needs of children with disabilities (including evaluation of individualized education programs)..." [20 U.S.C. 1413(a)(11)]. In addition, Section 435(b)(4) of the General Education Provisions Act calls for the State to evaluate the effectiveness of covered programs (including the IDEA) at least every three years.

Regulations in effect for both these laws provide essentially no guidance or specificity as to *how* evaluations are to be conducted. The only specific Federal requirements for evaluating special education programs, as expressed in the IDEA, are that it **be done annually, address program effectiveness**, and **include individualized education programs**. At this point in time, Federal monitoring activities surrounding the annual evaluation requirement are focused on the content of the State Plan (G. Hawkins, personal communication, July 12, 1993). Within the Plan, there must be a clear distinction between evaluation for effectiveness and compliance monitoring. Additionally, annual evaluation procedures or processes must be described, including a method for evaluating the IEP. There is no extensive review of the annual evaluation requirement during Federal on-site monitoring.

A recent study by the National Association of State Directors of Special Education (NASDSE) (Gonzalez, 1992) looked at State procedures for the annual evaluation of special education program effectiveness described in State Plans covering the 1992 school year. Also examined were the results of these evaluations found in States' IDEA Part B Performance Reports for 1989-90. The results of the study indicated a wide variety of evaluation activities among the States including short-term studies to evaluate specific issues, projects, or programs; statewide sampling and analysis of IEPs; and statewide evaluations of post-school (transition) outcomes or student performance outcomes.

The finding of broadest significance in the NASDSE study was that the majority of States (94 percent) used compliance monitoring procedures as the predominant program evaluation tool.

In some States, there were clear implications that the findings from compliance monitoring were used as evidence of the effectiveness of special education programs and services, and that corrective action following compliance monitoring was equated with implementing program improvements. Thus, based on the documents reviewed for that study, it appears that program evaluation and compliance monitoring tend to be viewed by State Education Agencies (SEAs) as the same activity.

This paper discusses four sub-topics related to program evaluation of special education. Discussion will focus on (a) suggested linkages and distinctions between effectiveness evaluation and implementation evaluation; (b) linkages between those two approaches, compliance monitoring, and the concepts of program improvement and accountability; (c) utilization of evaluation results; and (d) recommendations for a reconceptualization of the annual evaluation requirement under Part B of IDEA.

The Big Picture: Evaluating Effectiveness and Implementation

Implications of Effectiveness Evaluation

Summative evaluation is a conclusion-oriented approach which can be used to determine the effectiveness of a program. Patton (1986) tells us that "summative evaluation questions are these: Has the program been effective? [For whom?--Under what circumstances?--In what ways?] Should it be continued? Is it worthwhile? Did the program bring about (cause) the desired outcomes?" (p. 66). The measurement of outcomes is central to determining effectiveness and can be measured at the program or individual level. Outcomes evaluation essentially compares actual outcomes of a program with desired outcomes (goals). The measurement of student outcomes, in particular, is an important part of recent education reform initiatives in the states. The NASDSE study (Gonzalez, 1992) found numerous states (e.g., Connecticut, Delaware, Kentucky, Maryland, Michigan, Utah, Vermont, and West Virginia) in various stages of implementing statewide requirements for collecting student outcome information as a means of judging the effectiveness of school programs. Using student outcomes as a way of holding educators accountable for their services is gaining support nationwide[1].

---

[1] For persons who provide educational services to children with disabilities under Part B of IDEA, the "accountability" factor is explained in the Note that follows Section 34 C.F.R. §300.350: "This section is intended to relieve concerns that the IEP constitutes a guarantee by the public agency and the teacher that the child will progress at a specified rate. However, this section does not relieve agencies and teachers from making good faith efforts to assist the child in achieving the goals and objectives listed in the IEP. Further, the section does not limit a parent's right to complain and ask for revisions of the child's program, or to invoke due process procedures, if the parent feels that these efforts are not being made."

DeStefano and Wagner (1990) define outcome assessment as a focus on individual achievements, statuses, or behaviors.

> Assessments of such special education outcomes most often have an evaluative purpose in that outcomes measured for special education students (or former students) reflect how well the special education system in general is doing. Since most students with disabilities receive part of their education in the mainstream, outcome assessment can also describe and evaluate regular education programs. (p. 3)

Historically, outcome assessment has employed primarily survey methodology, sampling designs, and comparative analyses, particularly when the focus is on determining effectiveness. As DeStefano and Wagner state, often the underlying issue in judging program effectiveness using outcomes is "more effective than what?" Answering this question requires a comparison of student data to baseline levels, to established targets or standards, or to data from a comparison group.

The answers to questions concerning the effectiveness of a program are often useful in determining whether or not to continue to provide resources in support of it. Effectiveness evaluation, therefore, is often the method of choice for funding entities such as Congress, the Office of Special Education Programs (OSEP), the SEA, and state legislatures who must in some way account for the expenditure of public monies. In a real sense, effectiveness evaluation as a conclusion-oriented approach provides them with the direction they need - fund, don't fund; increase or decrease funding; fund some form of assistance or amelioration. In order to make decisions that will ultimately improve programs; however, program staff need the action-oriented information from implementation evaluation.

Implications of Implementation Evaluation

"Implementation evaluation focuses on finding out if the program has all its parts, if the parts are functional, and if the program is operating as it's suppose to be operating" (Patton, 1986, p. 124). Consequently, it can be used to make decisions on how to improve the implementation of a program. In its most basic form implementation evaluation can provide a historical written description of the program for dissemination or replication purposes. Perhaps most importantly, it provides crucial information to decisionmakers on whether or not a program is implemented according to design.

> When outcomes are evaluated without knowledge of implementation, the results seldom provide a direction for action because the decisionmaker lacks information about what produced the observed outcomes (or lack of outcomes). This is the "black box" approach to evaluation: Clients are tested before entering the program

and after completing the program, while what happens in between is a black box. (Patton, 1986, p. 129)

Imagine, for example, that a new curriculum is tested in one school district for possible use throughout the State. Student performance measures taken before and after exposure to the new curriculum indicate that performance improved, but there is no indication as to whether the curriculum was actually implemented as designed. Suppose, in fact, the teachers decided to use some of their own instructional materials in addition to those outlined in the curriculum. For some stakeholders it is perhaps enough to know that the new curriculum worked. The SEA, however, intends to introduce the curriculum across all districts, and unfortunately, they will do so without the knowledge that the curriculum was changed, perhaps for the better. These replications will not include all the components contributing to success in the pilot project.

The example above indicates the significance of implementation evaluation to an important group of stakeholders, those interested in dissemination and replication. In this case the purpose of the evaluation (to make a policy and programmatic decisions regarding the curriculum) would have been enhanced by the use of implementation evaluation in addition to outcome evaluation. There are a number of different types of implementation evaluation which can be used alone or in combination. For the purposes of this paper three types of implementation evaluation will be briefly introduced: program or compliance monitoring, process evaluation, and component evaluation.

Compliance monitoring is a form of implementation evaluation. It is one means of collecting data to answer questions such as, "Was the program implemented as designed"? or "Were the mandated guidelines and requirements being met"? As most SEA systems of compliance monitoring in special education have been devised, they rely heavily on the use of checklists to indicate the presence or absence of mandated process components; thus, it is one of the few strategies employed in implementation evaluation that can be summarized quantitatively (e.g., the percentage of completed IEPs, the number of students participating in transition planning). Being able to quantify the results of evaluation **does** allow the SEA to identify patterns of compliance within and across districts, including the areas of strength as well as the program components needing technical assistance. However, restricting evaluation to collecting quantitative data fails to provide a description of contextual variables, processes, or individuals influencing the provision of special education services. Without this information, it is difficult to determine where and how to make improvements and to understand the context necessary for program replication.

Process evaluation focuses on why certain things are happening, how the parts of the program fit together, and how people perceive the program as a whole (Patton, 1986). It is often a descriptive endeavor that documents "everything associated with *but not caused by* [italics added] the [program being evaluated]" (Scriven, 1991, p. 207). Its principal function is to

provide descriptive detail on the process of implementing a program so that defects in its design can be detected and corrected.

Component evaluation uses essentially the same naturalistic methods (e.g., interviews, participant observation) as process evaluation, however, the target of the evaluation is the operation of a specific component of the program rather than the entire program. This type of evaluation also can be used to study how various components are combined in a "complete" program evaluation (Scriven, 1991). Component evaluation would seem to be a particularly useful approach to evaluating special education as the various components (e.g., the IEP) are themselves rather complex processes that may be replicated or targeted for change as separate entities. Additionally, nearly all the processes in special education are linked in some manner; for example, student assessment for eligibility in special education is clearly affected to some degree by the referral process. In turn, the assessment process is directly linked to the IEP process. Scriven (1991, p.85) suggests that "since [program] components are frequently of variable quality, and since we are frequently looking for diagnoses that will lead to improvement, evaluating the components is sometimes a useful approach...how helpful depends upon the extent to which the 'fixes' for the defective components are self-evident or easily determined."

The types of evaluation outlined above, effectiveness and implementation evaluation, provide distinctly different information about a program; therefore, the relative emphasis of one type of evaluation over another or the benefits of combining both types in an evaluation system must be cautiously weighed. The overall purpose of the evaluation and the use of the results need to be considered in the design of a system of program evaluation in special education.

Defining the Purpose of Evaluation

Thorough consideration of the purpose(s) of, or reason(s) behind, program evaluation is a preliminary step to selecting the type of evaluation to be conducted, the evaluation questions to be posed, and the methodology to be employed. Two frequently cited reasons for evaluation are program improvement and accountability. These purposes have distinctly different implications for the conduct of program evaluation.

A System for Program Improvement

Of particular interest to many stakeholders is the notion of using evaluation results to improve programs, also known as the formative approach to evaluation. Questions posed in evaluation for this purpose include: What are the strengths and weaknesses of the program? How can the program be improved? What is working well and what is not working so well? What are the reactions of clients, staff, and others to the program? What are their perceptions about what should be changed? (Patton, 1986, p. 66) These are fundamentally different questions

from those posed in an effectiveness/outcomes approach to evaluation and generally require descriptive or naturalistic methods such as those used in process and component evaluation.

The annual evaluation requirement of the IDEA specifically includes a focus on the IEP. Although evaluating the effectiveness of the IEP is central to evaluating the effectiveness of special education, the actual processes for developing and using the IEP lend themselves to formative evaluation using descriptive methods. A naturalistic approach to evaluating the IEP might include, for example, observing IEP meetings to characterize how assessment information is used in the planning process or how goals and objectives are formulated or reformulated over time. A broad formative evaluation also could use case study methodology to explore the effects of the IEP process on individual participants.

Because of limited descriptive information, there is much we do not know about the IEP **process** itself and how best to configure it so that it becomes a useful mechanism for planning special education and related services for children and youth with disabilities. Available information on the IEP process has been derived either from the results of compliance monitoring, such as the presence/absence of short term objectives and completion according to timeline; or from SEA attempts to evaluate the outcomes of the IEP process using data on the number of goals or objectives completed (Gonzalez, 1992). The latter approach, is particularly prone to manipulation. Outcome data can be controlled by writing easily-obtainable goals and objectives, a definite problem if entities are to be held "accountable". Similarly, there is no means for evaluating the appropriateness of selected goals or objectives. Significantly, both outcomes-based accountability and compliance monitoring (as it is currently practiced) tell us more about the IEP *document* than the *processes* conducted to develop or use it. This is a serious concern since information on how the IEP process was developed lends credibility and validity to the document itself. Moreover, knowing how the contents of the plan actually were **used** in the provision of direct service, documents one of the underlying tenants of the statute.

A System for Accountability

According to Olsen and Massanari (1991) accountability, as a basis for evaluation, involves judging the degree to which a program was implemented as planned and produced the desired effects. Accountability is linked to both compliance monitoring (Is the program implemented as required?) and the evaluation of program effectiveness (What are the outcomes of the program, planned and unplanned?). It is important to note, however, that while compliance monitoring and evaluation both contribute to accountability, they are not the same thing (DeStefano, 1992).

Evaluating the outcomes of a particular program to determine its effectiveness is an important part of evaluation, but sometimes it is not enough to know that a program "worked". Often stakeholders also want to know whether—or to what extent—the program's most

important elements actually are being implemented (King, Morris, & Fitz-Gibbon, 1987). Thus, evaluation for the purpose of accountability profits from an implementation component - that component can be compliance monitoring. By way of illustration, due process is considered a critical element to guaranteeing the rights of children and families to a free and appropriate education. It is likely that advocates will continue to want to know the extent to which due process protections are provided in special education, independent of the achievement of positive outcomes for students with disabilities. Outcome evaluation alone will not address the ongoing implementation of this protection. In a sense, then, accountability in special education can have two aspects: procedural accountability and outcomes-based accountability (L. DeStefano, personal communication, October 1, 1993).

Porter (1993) discusses several methods for measuring school delivery standards[2] as part of an accountability system, including school site reviews for accreditation purposes. Using this "model", there would be no routine (cyclical) compliance monitoring in special education. Rather, a school would be judged based upon student performance on outcome measures. Schools whose student performance did not meet or exceed some criterion would be placed on probation. During probation, the schools **would be monitored** to assure that required processes and procedures were implemented as designed. Technical assistance would be provided to design and execute a corrective action plan.

Whether using this or another model, there is significant interest nationwide in devising a system for combining compliance monitoring and outcome evaluation for accountability purposes. There are, however, several potential problems associated with outcomes evaluation for accountability in particular that need to be considered. First, holding an entity accountable for their performance would seem to warrant a system of rewards and/or punishments (Porter, 1993). Any such system should, of course, be judged based on its effects on the overall goal of program improvement. Unfortunately, the use of monetary rewards or punishments are rife with potential problems not the least of which is that redistribution of resources away from poorly performing schools may hurt the children the most. After all, poor performance may result from a lack of will, a lack of know-how, or a lack of resources (Porter, 1993). Similarly, monetary rewards can encourage "cheating" to the extent that evaluation data can be manipulated to reflect better outcomes.

A second problem with outcomes evaluation in the context of accountability is that it requires criteria or a cutoff point between performance that is satisfactory and that which is not (Porter, 1993). Establishing criteria has been a long-standing problem in *procedural compliance* for both the SEA, and in some instances, the Federal government. Outcomes

---

[2] While there are a number of definitions for "school delivery standards", many include a designation of specific school inputs, required practices, or organizational characteristics. To some extent these items can be likened to the types of inputs, procedures or characteristics which are the focus of compliance monitoring in special education.

evaluation poses similar dilemmas: How many children can have poor outcomes before a school's performance is considered inadequate? Are certain outcomes more important to achieve than others? Are certain outcomes for certain children more important to achieve than other outcomes? The punishment, withholding or return of funds, is abundantly clear, but the absence of criteria makes accountability within the compliance monitoring system appear capricious at best. Perhaps part of the reluctance to establish firm criteria is the knowledge that sanctions are unlikely, in and of themselves, to result in program improvement.

A third potential problem in the use of evaluation data for accountability purposes is its effect on the participation of a major stakeholder in the evaluation - the local education agency (LFA). Aside from the general anxiety and animosity that accountability systems can engender, it is unrealistic to assume that LEAs could participate objectively in collecting data upon which their performance is judged. However, without the participation of local education agency (LEA) personnel, there is little hope that the results of the evaluation will be used at the district level.

Influencing the Utility of Evaluation Results

Thus far, this paper has only touched upon the impact of certain types of evaluation upon the use of results; however, a significant amount of research has been devoted to identifying types of evaluation use and significant contributing factors. For example, King and Pechman (1984) offer a dichotomous conceptualization of "use". In their scheme, one type of evaluation use is called "signalling" (Zucker, 1981), whereby the information provided through evaluation substantiates the successful completion of legal obligations. Evaluation for the purpose of accountability and, in most cases, compliance monitoring, serves a signalling function. As King and Pechman point out, the use of evaluation results as an accountability signal is clearly important for institutional survival, but has little bearing upon day-to-day practice. A quote from an LEA staff person in Evans and Weld (1989, p. 235) illustrates this point:

> I don't know what the State does with that information [Metropolitan Achievement Test Scores]. I don't know what we do with it. I don't think we do anything with it, other than send it to the State and then they send us money, which is not a bad trade.

Whether the Federal government, or in some cases the SEA, actually **uses** the evaluation data collected for accountability purposes is another issue (Suarez, 1990). As was mentioned earlier, the notion of accountability presumes a criteria for making judgements and a system of awards and/or punishments. Collecting data for collection sake—that is, when there is no apparent use of the data or consequences for what the data reveal—creates a perfunctory response on the part of those who collect it.

A second type of evaluation use, "charged use", carries with it the potential for disruption or change by giving individuals or groups of individuals data upon which to base action (King & Pechman, 1984). There are categories of charged use that more clearly describe the ways in which evaluation results can be utilized. The first category, *instrumental use*, occurs as a direct reaction to evaluation information. An example would be the discontinuation of a teacher recruitment program after evaluation showed it to be unsuccessful. A second type, *conceptual use*, transpires over time as the evaluation results influence a person's thinking about an issue. It may, in fact, be better to call this type of use "conceptual change". This category of use has likely occurred in connection with inclusion, as administrators, over time, have digested evaluative information and decided to adopt the practice. Lastly, when a user applies the evaluation process or results for personal ends (political or otherwise), this is referred to as *persuasive use*. Several examples of persuasive use occurred in the 1980s when evaluative information was used in some States to lobby for legislation regarding secondary transition services. Situational factors, such as political considerations, organizational variables, evaluator credibility, and resource constraints affect all forms of evaluation use (Alkin, Daillak, & White, 1979); however, the variable with perhaps the greatest impact on use is the extent to which major decisionmakers have been involved in the evaluation process.

In their review of the research on evaluation utilization, Cousins and Leithwood (1986, p. 360-361) conclude:

> Results argue strongly for evaluation procedures that at the outset generate information helpful to users in carrying out their decisions. Results also suggest that evaluation users should be involved in ways manageable for them, in the planning and carrying out of evaluation. Such involvement seems likely to ensure the credibility and relevance of the results, and to increase commitment to the evaluation process as a whole....

Decisionmaker involvement is a hallmark of the formative evaluation approach (e.g., implementation evaluation) and a necessary condition for program improvement.

Reconceptualizing the Annual Evaluation Requirement

This paper briefly described linkages and distinctions between two approaches to program evaluation: effectiveness and implementation, each of which can have distinctly different purposes, methods, and implications for use. In reviewing what these two approaches have to offer, clearly special education can benefit from both. The focus of the annual evaluation requirement within IDEA, Part B is on effectiveness (outcome) evaluation. While this does not preclude conducting other types of evaluation, the stipulation that *effectiveness* evaluation be done annually assumes that the majority of resources will be used for that purpose. **In implementing the evaluation requirement, SEAs need to determine the relative value of both information**

that permits a judgement of effectiveness and information that suggests improvements in implementation, particularly with respect to evaluating the IEP.

In addition, decisions at the Federal and State levels need to be made regarding the overall purpose of program evaluation in special education. This decision must be reached early in the design of the system since evaluation for institutional accountability can look very different from one based on program improvement, or from a system that addresses both purposes.

Another area in need of reconsideration is the marked imbalance between the resources devoted to compliance monitoring and program evaluation at the State and Federal levels. Monitoring systems are currently designed essentially to confirm or deny the existence of a particular program component. Recognizing this function, compliance monitoring has a small role in both effectiveness and implementation evaluation. If compliance monitoring determines that a component has not been implemented, there is essentially nothing in place to evaluate. Strategies to implement the missing component need to be devised. However, when monitoring establishes the existence of a program component, then there is a need for program evaluation to determine how it is configured and if it is effective. **It is important to note that the adjunctive role of monitoring in a reconceptualized evaluation system speaks strongly for reconsidering the extensive resources currently allocated to monitoring.**

Given the discrepancy between the resources demanded by any credible evaluation and those currently allocated to this activity, it is not feasible nor federally mandated to evaluate the effectiveness of all the components of special education (including the IEP) each year. States now have the flexibility to choose the focus of their annual evaluation, as long as the IEP is included. **Currently, states without an effectiveness and implementation evaluation system might look to approaches that outline an ongoing system of evaluation involving their LEAs. Through these approaches, the LEAs provide annual evaluative information on some aspect of special education. Such a system would allow evaluation resources to be concentrated on a few priority questions each year while the body of knowledge on both the effectiveness and implementation of special education programs and services accumulates over time. Other critical elements of an evaluation system, such as stakeholder input and the utilization of results, can be enhanced by a planned, progressive, long-term evaluation strategy implemented at the Federal, State and local levels.**

REFERENCES

Alkin, M.C., Daillak, R., & White, B. (1979). *Using Evaluations*. Beverley Hills, CA: Sage.

Cousins, J.B., & Leithwood, K.A. (1986). Current empirical research on evaluation utilization. *Review of Educational Research, 56*(3), 331-364.

DeStefano, L. (1992). Evaluating effectiveness: A comparison of federal expectations and local capabilities for evaluation among federally funded model demonstration programs. *Educational Evaluation and Policy Analysis, 14*(2), 157-168.

DeStefano, L., & Wagner, M. (1990). *Outcome assessment in special education: Lessons learned.* (Phase II, Deliverable 23, Contract No. 300-87-0155). Rockville, MD: Westat, Inc.

Evans, I.M., & Weld, E.M. (1989). Evaluating special education programs: Process and outcome. In D. Biklen, D. Ferguson, & A. Ford (Eds.), *Schooling and disability: Part II* (pp. 232-255). Chicago, IL: University of Chicago Press.

Gonzalez, P.A. (1992). *State procedures for the evaluation of special education program effectiveness*. Alexandria, VA: National Association of State Directors of Special Education, Inc., Project FORUM.

King, J.A., Morris, L.L., & Fitz-Gibbon, C.T. (1987). *How to assess program implementation.* Newbury Park, CA: Sage.

King, J.A., & Pechman, E.M. (1984). Pinning a wave to the shore: Conceptualizing evaluation use in school systems. *Educational Evaluation and Policy Analysis, 6*(3), 241-251.

Olsen, K.R., & Massanari, C.B. (1991). *Special education program evaluation: What should states consider?* Lexington, KY: University of Kentucky, Mid-South Regional Resource Center.

Patton, M.Q. (1986). *Utilization-focused evaluation* (2nd ed.). Newbury Park, CA: Sage.

Porter, A.C. (1993). School delivery standards. *Educational Researcher, 22*, 24-30.

Scriven, M. (1991). *Evaluation thesaurus*. Newbury Park, CA: Sage.

Suarez, T. (1990, October). *Living with the mixed message: The effects of government-sponsored evaluation requirements on the practice of evaluation.* Paper presented at the meeting of the American Evaluation Association, Washington, DC.

Zucker, L.G. (1981). Institutional structure and organizational processes: The role of evaluation units in schools. In A. Bank & R. C. Williams (Eds.), *Evaluation in school districts: Organizational perspectives* (CSE Monograph Series in Evaluation No. 10). Los Angeles, CA: University of California, Center for the Study of Evaluation.